**Information Visualization**

**PROJECT 3- Programming with D3.js**

**Team Members - 3:  
 Mukthasree Vengoti (811296123),**

**Keerthi Akhila Pasam (811304142),**

**Subhasmita Maharana (811263851)**

**Project Webpage link:** [**https://subhasmita23-s.github.io/Mutualfund-vis/**](https://subhasmita23-s.github.io/Mutualfund-vis/)

**Project GitHub link:** [**https://github.com/Subhasmita23-S/Mutualfund-vis**](https://github.com/Subhasmita23-S/Mutualfund-vis)

**Introduction:**

The goal of this project is to visualize mutual fund data in order to identify important patterns and trends that will help investors make well-informed decisions. We created three interactive visualizations using D3.js to investigate the distribution of assets within mutual fund categories, the link between expenditure ratios and fund returns, and fund performance over a range of time horizons. User-friendly interaction elements such as brushing, tooltips, zooming, filtering, and dynamic sorting improve each visualization. The solution provides a thorough web-based interface for financial insights and facilitates data-driven exploration using cleanly processed datasets.

**Visualization Tasks:**

**Task 1: Compare Fund Performance Over Time**

“Which funds have consistently performed well over 3MO, 1Y, 3Y, 5Y, 10Y, and YTD periods?”

**Objective of Task 1:** The objective of this job was to assess the performance of mutual funds across the following return periods: 5-year (5YR), 10-year (10YR), 1-year (1YR), 3-year (3YR), 3-month (3MO), and year-to-date (YTD). Helping investors choose funds that exhibit steady performance over both short- and long-term horizons was the main goal to promote more informed investment choices.

**Task 2: Analyze Relationship Between Expense Ratio and Total Assets**

“Is there a relationship between a fund’s expense ratio and its performance?”

**Objective of Task 2:** To determine whether higher management fees are warranted by long-term returns and whether there is a trend between cost and return potential, the task's objective was to investigate the relationship between a mutual fund's expense ratio and performance over a variety of time horizons (3MO, YTD, 1YR, 3YR, 5YR, and 10YR).

**Task 3: Explore Fund Categories Based on Return Distributions**

“How are mutual funds distributed across categories and how much in assets do they hold?”

**Objective of Task 3:** The objective of this job is to use a tree map to show the distribution of mutual fund assets among several fund categories. With size corresponding to total assets and color denoting relative magnitude, each block in the tree map represents a category. The user experience and readability are improved by interactions like zooming, tooltips, and reset.

**Visualization & Interaction System Designs:**

**Task 1: Compare Fund Performance Over Time**

**Visualization Design Representation: Grouped Bar-chart**

To efficiently display multi-year return data across many funds, a Grouped Bar Chart was used for this task. The X-axis shows each mutual fund, while the grouped bars show returns across the six chosen time periods.

This format was chosen because it enables:

* A clear visual comparison of performance over a few time periods.
* The ability to quickly identify volatile and high-performing funds.
* The ability to compare a configurable list of top funds simultaneously.

A graph of different colored lines

AI-generated content may be incorrect.

**Interaction Design Representation:**

**Problem Without Interaction:**

* Clutter: Hundreds of lines overlap and make it unreadable.
* Hard to focus: It's difficult to track one fund across multiple time periods.

**Why Needed?**

* Reduce clutter.
* Allow users to focus only on a subset of important funds.
* Dynamic exploration based on time horizon preferences.

The interactivity design below makes sure that even with enormous datasets, the chart is visible and useable without taking up too much space on the user interface.

|  |  |
| --- | --- |
| **Features** | **Purpose** |
| Brushing (Top10, Top20, Top50) | Allows users to view more or fewer funds based on sorted performance. |
| Sorting Dropdown (Sort by Return) | Sorts the displayed funds by selected time-period (e.g., 1YR, 10YR) |
| Labeling (Tooltips on Hover) | Shows fund names and exact return percentages for each bar on mouseover. |
| Horizontal Scroll | Ensures readability even when displaying many funds (like Top 50) |
| Hover Highlighting | Highlight a single fund's path when mouse hovers over it. |

**Task 2: Analyze Relationship Between Expense Ratio and Total Assets**

**Visualization Design Representation: Scatter plot**

To illustrate this link, a scatter plot was selected, with the expense ratio, expressed as a percentage of the assets under management, is the X-axis. Y-axis, the chosen time period's return percentage. One mutual fund is represented by each dot.

This design makes it possible to:

* Identify outliers for abnormally high fees or very high/low returns.
* Visually correlate cost and performance quickly.
* A dynamic comparison between many return periods.

**Interaction Design Representation:**

**Problem Without Interaction:**

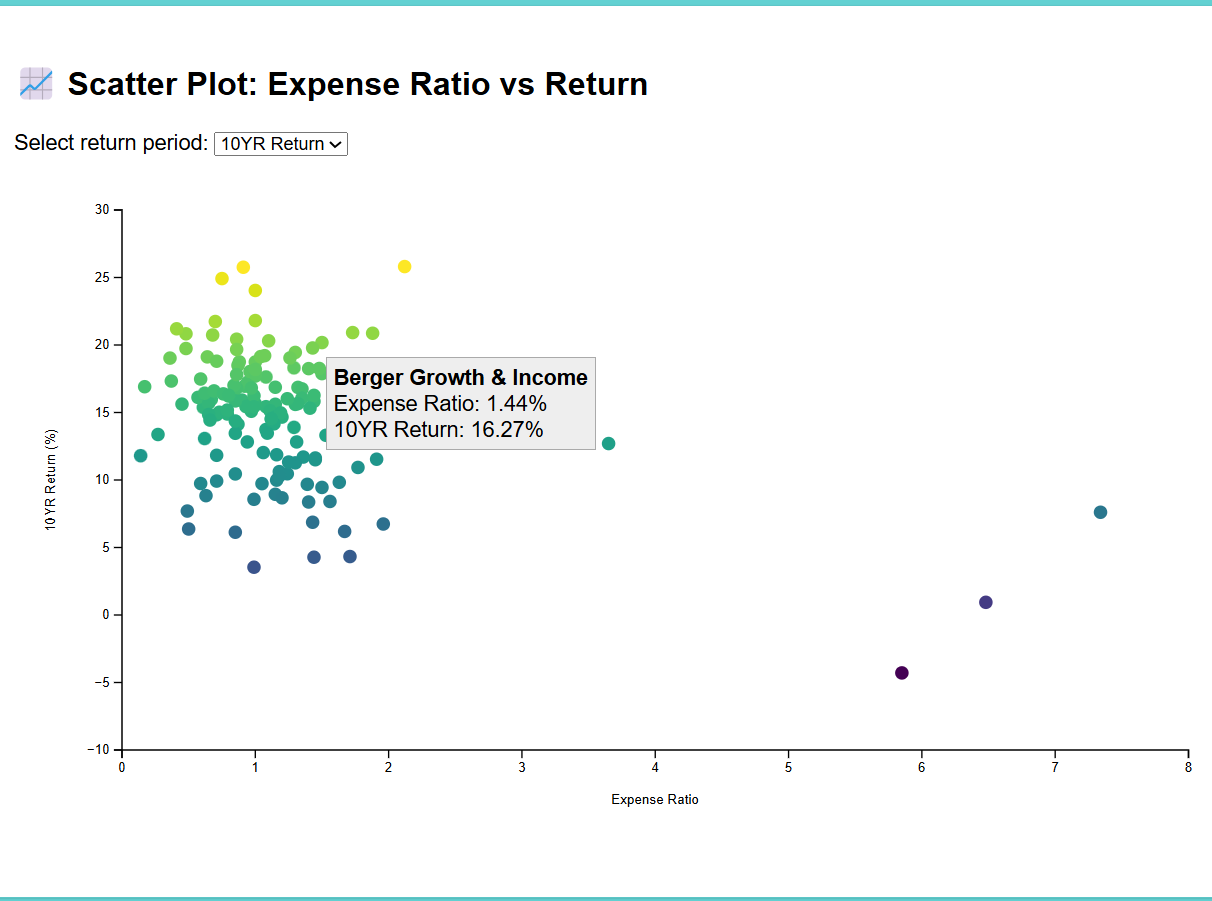
* Box plots show category-level information only (aggregate view).
* Users can’t see individual fund returns or exact median values easily.

**Why Needed?**

* Drill down from general category statistics to individual fund performance.
* Create an interconnected view between fund returns and categories.
* Help identify outlier funds easily.

The visualization supports the following interactive features:

|  |  |
| --- | --- |
| **Features** | **Purpose** |
| Dropdown (Return period) | Allows users to switch between 3MO, YTD, 1YR, 3YR, 5YR, and 10YR returns |
| Color Encoding (Viridis) | Enhances visual contrast based on return value (darker = lower, brighter = higher) |
| Dynamic Axes & Filtering | Automatically adjusts the Y-axis based on data distribution (excluding 0s and Nan’s) |
| Tooltip | Displays fund name, expense ratio, and exact return percentage on hover |



**Task 3: Explore Fund Categories Based on Return Distributions**

**Visualization Design Representation: Tree map**

Mutual Fund Categories are hierarchical (categories → funds). Tree maps allow us to show **parts of a whole** naturally. Tree maps make excellent use of limited screen space to show many categories. Area size encodes total assets:

* Larger rectangles = Categories with larger total assets
* Color can represent another dimension (e.g., Expense Ratio)

**Filtering Relevant Fields:**

We extracted only the required columns:

* Category
* Net Assets

**Data Cleaning:**

* Rows with missing or invalid values in Category or Net Assets were removed.
* Net Assets values were cleaned and converted to numeric format by stripping out currency symbols and commas.

**Grouping and Summarizing:**

We grouped the data by Category and applied the following aggregations:

* **Total Assets** = Sum of Net Assets for each category.
* **Fund Count** = Number of unique funds in each category.
* **Asset Share Percent** = (Total Assets / Sum of all assets) × 100

The resulting CSV file looks like this:

Category, Total Assets, Fund Count, Asset Share Percent  
Large Value,243467891250.0,232,41.15  
Large Growth,181107779592.0,196,30.61  
International Stocks,110835524547.0,279,18.73  
Small Growth,35731833916.0,156,6.04  
Small Value,20503116359.0,118,3.47  
This cleaned and aggregated data was then used for the D3.js-based tree map visualization.

**Interaction Design Representation:**

**Problem Without Interaction:**

* Small categories are difficult to read.
* Static tree map limits exploration of deeper categories or fund-level details.

**Why Needed?**

* Explore finer categories without clutter.
* Understand asset concentration visually.
* Make tree map blocks readable even for small categories.

|  |  |
| --- | --- |
| **Features** | **Purpose** |
| Block Size | Proportional to Total Assets. |
| Tooltip on Hover | Shows category name, total assets, number of funds, and asset share %. |
| Zoom-in on click | Click on any category block to zoom into it. |
| Reset Zoom Button | Allows users to reset the zoom and return to full view. |
| Color Scale | Fills are generated using d3.interpolateBlues based on relative asset size. |
| Label readability | Font size and text color dynamically adjust based on background brightness. |

**Web-based Visualization System:**

The Web-based Visualization System for this project is designed via standard browsers. It is included with visualization designs along with interaction features which are listed below:

**Technologies Used:**

|  |  |
| --- | --- |
| **Tools** | **Purpose** |
| HTML/CSS | Structure and styling of the webpages |
| Javascript | Scripting for logic and interactions |
| D3.js (v7) | Core library for data-driven visualizations |
| CSV files | Data format used for each task |
| GitHub pages | Web hosting platform for the live system |

**Implementation and Interaction Highlights:**

**Task 1: Compare Fund Performance Over Time**

* To create a parallel coordinate graphic, D3.js was used.
* The chart's lines, which go across the following axes, each represent a mutual fund:  
  YTD, 1YR, 3YR, 5YR, and 10YR accruals
* Data from task1\_mutualfunds\_cleaned.csv is loaded by the script (task1\_script.js).

**Dropdown menus:** Allow users to sort and filter funds according to return periods (YTD, 1YR, etc.). Permit the Top 10, 20, or 50 funds to be selected.

**Brushing & Highlighting:** A fund line is highlighted throughout the chart when you hover over it.

**Coordination:** The shared dispatcher (events.js) links Fund selection to Task 2.

**About the graph:** The performance of the top 50 mutual funds, arranged by 10-year returns, is shown in the bar chart. A particular mutual fund is represented by each group of bars, and the returns of the fund over the following time periods are shown by the different colored bars within each group: three months (3MO), year-to-date (YTD), one year (1YR), three years (3YR), five years (5YR), and ten years (10YR).

It uses the 10-year returns of the top 50 mutual funds to explain their past performance. A visual comparison of these top-performing mutual funds' past returns is made possible by the graphic. This enables you to observe these top funds' recent performance trends in addition to their long-term success.

A screen shot of a graph

AI-generated content may be incorrect.

**Task 2:** **Analyze Relationship Between Expense Ratio and Total Assets**

* Used D3 to create a box plot, one for every mutual fund category.
* Every box display:
* One-year return median, quartiles, and outliers
* Task2\_mutualfunds\_cleaned.csv was used to load the data.

**Tooltips on hover:** The precise fund return, and category details are displayed.

**Brushing & Coordination:** Highlighted fund from Task 1 appears in this chart.

**Color Mapping:** Optional color scale for categories or returns ranges.

**About the graph:** The scatter plot shows how different mutual funds' 5-year returns relate to their expenditure ratios. A single mutual fund is represented by each dot, whose vertical position displays the fund's 5-year return % and its horizontal position its expense ratio. Another quantity, like total assets, could be represented by the color of the dots; a gradient could suggest higher or lower asset values.

It describes how the expense ratio—a measure of a mutual fund's management costs—and the returns the fund has produced over a given time frame—currently set at five years—may or may not be correlated. You can determine whether there is a discernible pattern or whether funds with lower expense ratios often have higher returns by examining the distribution of the dots. Analysis of this relationship across various investment timelines is made possible by the flexibility to alter the return term.

A screen shot of a graph

AI-generated content may be incorrect.

**Task 3: Explore Fund Categories Based on Return Distributions**

* Built using D3’s tree map layout.
* The categories of mutual funds are represented by each rectangle, with:
* Amount in relation to Total Assets.
* Mapping the intensity of color to total assets.
* The task3\_mutualfunds\_cleaned.csv file is the source.

**Click-to-zoom:** Clicking on a category allows users to enlarge it for a deeper look.

**Reset Zoom**: Button to return to original full view.

**Tooltip on hover**: Displays fund category name, asset value, share, and fund count.

**Legend**: Color gradient to interpret total asset sizes visually.

**About the graph:** The mutual fund categories are shown in the tree map, and the size of each rectangle indicates the total amount of assets owned in that category. The categories with the highest total assets are shown by the largest rectangles. According to this perspective, "Large Value" and "Large Growth" seem to own a lot more assets than "International Stocks," "Small Growth," and "Small Value."

The distribution of total mutual fund assets across various investment categories is shown graphically in this tree map. The entire amount of money held in funds that fall under that category is closely correlated with the size of each colored rectangle. A multifaceted comparison of the relative importance and characteristics of different mutual fund segments is made possible by the color intensity, which may also help to describe a feature of these categories, possibly connected to their returns or risk levels.

A screenshot of a graph

AI-generated content may be incorrect.

**Personal Contributions:**

**Data Preparation:** Led the extraction, cleaning, and transformation of mutual fund data to prepare it for category-based analysis. This included handling missing values and formatting asset values.

**Visualization:** Designed and implemented the Tree Map using D3.js to represent mutual fund categories and their total asset distribution.

**Interactivity:** Added zoom-in, reset, and hover tooltip interactions to enhance user engagement and readability for smaller categories.

**Collaboration:** Assisted in coordinating the visualization system by aligning the format and functionality of the scatter plot with the rest of the dashboard.

**Conclusion:**

Through this project, we were able to effectively develop a browser-based, interactive visualization system that analyzes the performance and attributes of mutual funds using D3.js. By monitoring consistent performers, analyzing cost-return correlations, and comparing asset distributions at the category level, the three activities offered a comprehensive perspective. Thanks to user-friendly interfaces and organized layouts, users may now effortlessly examine intricate financial data visually. In order to improve decision-making, this research shows how information visualization can make large-scale investment data both understandable and accessible.